Identification and characterization of tebuconazole transformation products in soil by combining suspect screening and molecular typology

Veronika Storck, Luigi Lucini, Laure Mamy, Federico Ferrari, Evangelina Papadopoulou, Sofia Nikolaki, Panagiotis Karas, Rémi Servien, Dimitrios Karpouzas, Marco Trevisan, et al.

Identification and characterization of tebuconazole transformation products in soil by combining suspect screening and molecular typology

Veronika Storck[1,2,3], Luigi Lucini[3]*, Laure Mamy[4], Federico Ferrari[2], Evangelia S. Papadopoulou[5], Sofia Nikolaki[6], Panagiotis A. Karas[5], Remi Servien[7], Dimitrios G. Karpouzas[5], Marco Trevisan[3], Pierre Benoit[4], Fabrice Martin-Laurent[1]*

[1]INRA, UMR 1347 Agroécologie, Dijon, France

[2]Aeiforia srl, Spinoff Università Cattolica del Sacro Cuore, Fidenza, Italy

[3]Università Cattolica del Sacro Cuore, Instituto di Chimica Agraria ed Ambientale, Piacenza, Italy

[4]INRA, UMR 1402 ECOSYS, Thiverval-Grignon, France

[5]University of Thessaly, Department of Biochemistry and Biotechnology, Larissa, Greece

[6]University of Patras, Department of Environmental and Natural Resources Management, Agrinio, Greece

[7]INRA, UMR 1331 Toxalim, Toulouse, France

*Corresponding authors: fmartin@dijon.inra.fr; luigi.lucini@unicatt.it

Keywords: time-of-flight mass spectrometry, pesticide, triazole, clustering, environmental risk assessment.

ABSTRACT

Once released into the environment, pesticides generate transformation products (TPs) which may be of (eco-)toxicological importance. Past studies have demonstrated the difficulty to predict pesticide TP occurrence and their environmental risk by monitoring-driven approaches mostly used in current regulatory frameworks targeting only known toxicologically relevant TPs. We present a novel combined approach which identifies and categorizes known and unknown pesticide TPs in soil by combining suspect screening time-of-flight mass spectrometry with *in silico* molecular typology. This approach applies an empirical and theoretical pesticide TP library for compound identification by both non-target and target time-of-flight (tandem) mass spectrometry and structural elucidation through a molecular structure correlation program. *In silico* molecular typology was then used to group the detected TPs according to common molecular descriptors and to indirectly elucidate their environmental properties by analogy to known pesticide compounds having similar molecular descriptors. This approach was evaluated via the identification of TPs of the triazole fungicide tebuconazole occurring in a field dissipation study. Overall, 22 empirical and 12 yet unknown TPs were detected and categorized into three groups with defined environmental properties. This approach combining suspect screening time-of-flight mass spectrometry with molecular typology could be extended to other organic pollutants and used to rationalize the choice of TPs to be intensively studied towards a more comprehensive environmental risk assessment scheme.

CAPSULE ABSTRACT

Combination of time-of-flight mass spectrometry and molecular typology was shown to be of fundamental interest to identify and categorize tebuconazole transformation products in soil.

INTRODUCTION

Pesticides, used in agriculture to ensure crop yield and quality, constitute a large group of xenobiotic compounds in the environment.[1,2] They and their transformation products (TPs) have globally been identified as major contaminants of natural water resources. The United States Environmental Protection Agency (U.S. EPA) estimates a global release of 2.4 billion kilograms pesticides per year supporting important economic sectors of agriculture but also causing major environmental pollution.[3] Another aspect which makes pesticides a chemical group of particular interest is their rapidly changing and evolving market due to the continuous introduction of new molecules. In order to reach the market, new pesticide molecules have to undergo an environmental risk assessment (ERA) procedure based on a large set of data submitted by the pesticide companies and evaluated by the official authorities (e.g. the Plant Protection Products and their Residues panel[4] of the European Food Safety Authority for Europe (EFSA)). This procedure is periodically repeated in order to re-evaluate the already registered pesticide molecules. Once released into the agro-ecosystem, pesticides and their TPs can diffuse away from their target point, contaminate surrounding water resources[5] and impact non-target organisms supporting key ecosystemic services.[6] The properties of TPs such as bioavailability, mobility and ecotoxicity can be substantially different from the parent compound[7,8] and although their biological efficacy is usually lower,[9] they could be highly (eco-)toxicologically relevant.[10] In Europe, *a posteriori* monitoring of the contamination of water and air resources by pesticides is imposed by corresponding directives.[11,12] In contrast, no monitoring of soil contamination by pesticides is required at EU level. This is the result of the lack of a corresponding directive, although the need for a regulation was proposed to the European Commission.[13] Thus, if an unacceptable risk for the environment becomes evident several years after the introduction of a pesticide to the market, restrictions in its use are imposed and if monitoring studies reveal that no alleviation of the

risk was achieved, the pesticide is taken off the market. Past experience has shown that the knowledge of the existence of (eco-)toxicologically relevant TPs typically emerges only about 20 to 30 years after the introduction of a given pesticide to the market.[10] Currently used monitoring-driven approaches for pinpointing environmentally and (eco-)toxicologically relevant pesticide TPs often concentrate on water samples and do not prevent the contamination of natural resources and impose mitigation actions only after contamination is reported. Novel approaches able to identify potential TPs and *in silico* assessment of their toxicity and environmental contamination could be part of the solution of this issue.

The gap of knowledge about the existence or properties of pesticide TPs derived by abiotic and biotic processes in the environment may be the reason why relevant TPs have been overlooked for so long. Conventional analytical methods focus on the monitoring of the so-defined '(eco-)toxicologically relevant' pesticide TPs or TPs 'relevant for groundwater contamination'.[4,14] Consequently, most studies focus on a limited number of molecules which have been identified as targets by monitoring-driven approaches and are calibrated against analytical standards. Therefore, they often provide only a partial view of the complex transformation patterns of pesticides in the environment. Thus, there is a need to develop an innovative approach by combining powerful analytical methods detecting also unknown TPs with categorization methods based on the chemical features of the detected TPs.

The analytical techniques suitable to detect unknown and suspected molecules in the absence of analytical standards are high resolution mass spectrometry (MS), namely quadrupole time-of-flight MS (QTOF-MS) and Fourier transform MS (FT-MS), as well as nuclear magnetic resonance (NMR) spectroscopy.[15] These techniques are often combined in forensics[16] or medicine,[17] where QTOF-MS is used for screening for new unknown molecules and then is supplemented by NMR spectra for the structural elucidation of a compound. However, NMR spectroscopy is an effective method for the confirmation of molecular structures but is rather

less appropriate for the analysis of TPs in environmental samples due to its rather poor sensitivity, which generally is three orders of magnitude lower than MS.[15]

Categorization of unknown and suspected molecules can be achieved by *in silico* molecular typology that enables the property-related clustering of potential TPs.[18]

In this study, we present a new approach for the *a posteriori* analysis of agricultural soil samples treated with a pesticide by combining high-resolution QTOF-MS with *in silico* molecular typology. Our approach enables profiling and molecular structure elucidation of unknown and suspected pesticide TPs in agricultural soil by the use of a compound library containing empirical (already known) and theoretical (yet unknown) TPs. The proposed approach consists of three principal steps: (i) construction of an in-house pesticide-specific TP library based on literature survey and establishment of suspect TPs, (ii) QTOF-MS analyses of soil extracts by non-target MS to screen TPs against this library, followed by target tandem MS using an ion inclusion list to propose their structure via accurate mass daughter ions and a molecular structure correlation (MSC) program and (iii) categorization of detected TPs based on *in silico* analysis using molecular typology.

To demonstrate the potential of this novel approach, the dissipation of the triazole fungicide tebuconazole (TCZ) was investigated in a field dissipation study. TCZ was chosen as a model compound because, despite being on the market for quite some time (27 years),[19] there are still significant knowledge gaps regarding its environmental fate, its transformation in soil and the ecotoxicological impact of potential TPs.[20] The potential formation of various triazole TPs known to be recalcitrant to biodegradation and to possibly interact with the hormone regulation of non-target organisms by inhibition of the cytochrome P450-dependent conversion of lanosterol to ergosterol is particularly relevant.[21,22] The applied approach is of interest for improving future environmental fate studies to become more comprehensive, thus strengthening ERA.

MATERIAL AND METHODS

Library setup of TCZ TPs

An in-house TCZ TP library was constructed based on a thorough literature review including research papers (cited in the results section entitled 'Empirical TCZ TPs') and data sets of the pesticide company on the transformation pattern of TCZ in various matrices. Search was done by means of 'Web of Science', 'Google' and others using search terms like 'tebuconazole', 'metabolite', 'transformation product', 'degradation' and others. The developed library was composed of 47 empirical (already detected in at least one study) and 29 theoretical TPs. The latter were created based on expert knowledge taking into account common reaction processes in organic chemistry and biochemistry likely leading to probable molecules. The TCZ TP library contains the name, molecular formula, molecular structure and mass of each molecule (Supporting Information, table S1). Initial pretests of chemical analysis were carried out on a subset of soil samples randomly chosen from the field study to search for TCZ TPs. Retention times of TPs proposed via MSC were then added to the library as an additional parameter for identification purposes in MS screening.

TCZ field dissipation study

The field dissipation study was conducted on a North Italian field (45°04'80''N, 09°45'60''E) in autumn/winter 2013/14. The field soil was characterized as loamy sand (4.2 % clay, 13.5 % silt, 82.3 % sand) with a pH of 7.3 and 4.5 % of organic matter. It was not treated with pesticides for more than five years and cropped with winter wheat. A commercial formulation of TCZ (Folicur SE, Bayer Garden) was applied on field plots of 60 m^2 each at 2.5 kg TCZ/ha (5x the recommended dose).[23] Soil samples were taken at 0, 3, 7, 21, 35, 56, 70, 105 and 125 days after pesticide application according to the ISO standard.[24] For each plot and each sampling date, a composite soil sample was prepared from 10 sub-samples collected with an iron cylinder of 5 cm diameter, taking a soil depth of 5 cm. Soil samples were manually

148 homogenized, sieved (5 mm) and kept frozen at -20°C until further processing. The field
149 dissipation study was part of a larger study investigating field dissipation of three model
150 pesticides including TCZ at various application rates (four plots are used in this study: two
151 non-treated control plots and two TCZ 5x dose plots). The field dissipation of TCZ was
152 monitored in soil extracts via high performance liquid chromatography coupled to photodiode
153 array detector (HPLC-PDA) and followed a biphasic pattern with a half-life (DT_{50}) of 2.5
154 days, as calculated with a Hockey-stick kinetic model.[25] After 125 days, 14.2 % of applied
155 TCZ were remaining in the investigated soil layer.

156 **TCZ and TP extraction from soil**

157 TCZ and its related TPs were extracted with acetone from 40 g of soil as explained in the
158 Supporting Information (SI-2). The final obtained residues were re-dissolved in 1 mL of
159 70 v% methanol and 30 v% ultrapure water containing 0.01 % phosphoric acid and used for
160 chemical analysis. The extraction efficiency for TCZ was 86.36 % ± 0.45 %.

161 **UHPLC-ESI-QTOF MS analysis**

162 In order to investigate TCZ TPs, samples were first analyzed by non-target UHPLC-ESI-
163 QTOF mass spectrometry. TCZ TPs were primarily identified by screening of the MS-only
164 raw data against the TP library, using accurate mass, isotope spacing and isotope ratio for
165 identification. A difference in accurate mass of 5 ppm, together with an isotopic pattern
166 overall score higher than 85 % were used as cutoff values. Compounds that passed both
167 thresholds, that had a plausible peak chromatogram feature and that were not detected in the
168 controls (non-treated soil samples), were chosen for further analysis by target UHPLC-ESI-
169 QTOF tandem MS. For this purpose, the three most abundant precursor ions (generated
170 during MS-only) and the retention time (RT) of each MS-only detected TP were exported to
171 be included into a target list for tandem MS. Method details are explained in the Supporting
172 Information (SI-3).

Molecular structure correlation

Elucidation of the molecular structure of compounds detected by tandem MS was obtained through a molecular structure correlation (MSC) program (Agilent MassHunter Molecular Structure Correlator B.05.00). The software correlated accurate mass daughter ions for a compound of interest with the molecular structures available in the in-house custom library by using a 'systematic bond-breaking' approach as described by Hill and Mortishire-Smith.[26] Briefly, an overall correlation score was generated by taking into account the individual scores from each daughter ion, the mass accuracy of these fragments, and the overall percentage of ion intensities being plausibly explained with substructures.

Molecular typology for categorization of detected TCZ TPs

TCZ and its detected TPs were categorized using the molecular typology approach 'TyPol' (Typology of Pollutants) proposed by Servien et al. (2014) allowing clustering of chemicals based on the use of molecular descriptors and of environmental parameters.[18] Environmental parameters of TPs of TCZ being unknown, clustering was only done using a set of 40 molecular descriptors. After clustering, TyPol was used to estimate the environmental parameters (vapor pressure, dissipation half-life, sorption coefficient and bioconcentration factor) of TCZ TP clusters by analogy to reference compounds for which environmental parameters are known. Details about the method are given in the Supporting Information (SI-4).

RESULTS AND DISCUSSION

Empirical TCZ TPs

Literature review revealed that studies reporting TCZ dissipation were often ignoring potential TP occurrence and were mainly following targeted analytical approaches.[27,28] It is noteworthy that newly discovered TPs detected by non-target approaches were observed during degradation of TCZ by bacterial or fungal cultures,[29] as photodegradation products,[30]

198 in field soil studies[31,3], in soil microcosm studies[33,34] and in fruits.[35] Additional information
199 was found in regulatory documents mainly citing unpublished or confidential studies
200 reporting TPs in soil and crops partially based on MS and NMR spectroscopy.[36,37,38]
201 Altogether, literature review led to the construction of a library containing 47 empirical and
202 29 theoretical TCZ TPs (Supporting Information, table S1) which was used for TCZ TP
203 suspect screening in soil samples from the field dissipation study described before.

204 **Detection of TCZ TPs**

205 The QTOF-MS analyses of soil extracts gained from the field dissipation study led to the
206 detection of the parent compound TCZ, as well as of 22 empirical and 12 yet unknown TPs,
207 which were contained in the library. Among the 34 detected features, some of them showed
208 the same exact mass and retention times. 12 TPs passed all steps of identification from non-
209 target MS-only identification to target tandem MS and structure correlation. Three TPs passed
210 only the first two steps (selected by tandem MS but not further identified through structural
211 correlation) and 19 TPs were only detected by MS identification with plausible chromatogram
212 features, but were not further selected by tandem MS due to either method restrictions or used
213 selection criteria or molecular characteristics (e.g. low mass, high volatility,…). Table 1
214 summarizes the data of all detected TCZ TPs, reporting compound name, molecular structure,
215 mass, reference if empirical, and retention time of each TP. Moreover, the time points are
216 provided (expressed in days after the application of the TCZ treatment) at which the TCZ TPs
217 were detected and information regarding the confidence of identification. TPs are coded
218 according to their level of identification confidence (a and orange hooklet: detected by MS-
219 only; a,b and yellow hooklet: detected by MS and selected by tandem MS; a,b,c and green
220 hooklet: detected by MS, selected by tandem MS and identified by MSC). According to
221 Schymanski et al. (2014) level ‘a’ (MS-only detection) could be assigned to ‘level 4 -
222 identification confidence by unequivocal molecular formula’; level ‘a,b’ (selection by tandem

MS) to ‘level 3 - tentative candidate - identification confidence by structure, substituent and class’; level ‘a,b,c’ (structure proposition by MSC) to ‘level 2,b - probable structure - identification confidence by diagnostic evidence’. [39] The molecular formulas of compounds that have only been detected by non-target MS but were not further selected by target tandem MS are rather unlikely to be false positives, as they (i) were not detected in the controls, (ii) have plausible chromatogram features and isotopic patterns and (iii) do not differ in RT between replicates and different time points. However, their molecular structure rests tentative as the compound was not selected by tandem MS and as the molecular structure elucidation by MSC is only possible based on the detected product ions produced during tandem MS. More likely, reasons for the impossibility of selection by target tandem MS could be either peak saturation at high concentrations (as it has been seen for the parent molecule TCZ when injected at high volumes), or a high sensitivity of compounds present at low concentrations, or too low masses of the product ions or a high volatility of the precursor. Indeed, electron spray ionization (ESI) conditions and the chosen collision energies have to be efficient enough to produce precursor and product ions while avoiding fragmenting the precursor into too small product ions, as the minimum m/z value was chosen to be 60 due to noise exclusion. Regarding low m/z values, it must be mentioned that the low molecular weight of putative TPs not detected by the non-target MS-only screening such as the well-known triazole fungicide TP 1-H-1,2,4-triazole (TP 63)[35-38] may be considered as false negatives as they may not be efficiently ionized by ESI due to their volatility at the high temperature of the used ion source. Nevertheless, it should be noted that a compound, which was detected by MS-only and selected by target tandem MS but whose molecular structure could not be proposed by MSC, might be correct in its molecular formula but might have a molecular structure not present in the TP library. Additionally, the poor observed correlation for the low molecular weight TPs might be affected by the formation of product ions, which were below the m/z

threshold and thus were missing for MSC. Strictly, all the detected molecules by our approach still remain inferred until confirmed by retention time and fragmentation pattern of pure standards, which are not available on the market for most of them.

Despite of this weakness, the benefits of suspect screening coupled to MSC for the elucidation of molecular structures represent a powerful alternative when standards are not available. These benefits can be illustrated by the example of TPs 13 and 15, which are constitutional isomers. As they have the same molecular formula and mass, they could not be discriminated by MS-only and tandem MS (Supporting Information, figure S1) even using MS instruments with higher resolution power. MSC was able to correlate the product ions to the respective precursor for elucidation of their molecular structure (Supporting Information, figure S2) and hence had the capacity to discriminate between these TPs without need of any additional separation step or the use of standards. Furthermore, it must be pointed out that an even higher number of TPs can be taken into account through suspect screening than in targeted approaches, thus providing more comprehensive information on the transformation of one or more pesticides to a series of TPs in the environment.

The pesticide dissipation study carried out at field scale was found to be suitable for the purposes of our study. The applied TCZ concentration was in the range of concentrations commonly applied in agriculture[23] and nine time points within 125 days after treatment enabled the comprehensive exploration of TP formation and decay patterns although detailed data on the temporal patterns were not established due to the absence of analytical standards. As TCZ TPs were extracted from the top 5 cm layer of soil, it has to be noted that only those TPs present at this soil layer could be studied. TP properties can strongly deviate from those of the parent molecule[40] and might be vulnerable to leaching, runoff or volatilization although the parent molecule is considered as relatively immobile or non-volatile.[36,41] Thus, TPs which potentially dissipated from the top soil layer, which were not extractable from soil via the

applied extraction method or which were not listed in the TP library were disregarded in this study.

Moreover, TCZ and several of its TPs have two enantiomers or do even possess more than one chiral center. This may be important concerning stereoselective toxicity and dissipation[42] but it was not considered in this study. For the completion of our analytical approach and for the provision of comprehensive information on actual environmental fate, molecular typology can be applied to estimate potential properties of the detected TPs in a theoretical manner. This technique can be used to estimate properties of TPs of any anthropological substance in nature.

Categorization of TCZ TPs

TCZ and its 34 detected TPs were categorized by TyPol into 3 different clusters (table 1, figure 1) according to their molecular descriptors including molecular weight, number of atoms, dipole moment, polarizability or total energy (Supporting Information, table S2). TCZ and 23 of its TPs were categorized into cluster 1 exhibiting the highest simple and valence molecular connectivity indices (MCIs), energy of the highest occupied molecular orbital (E_{HOMO}), molecular weight (MW), surface (S), polarizability (α), number of atoms and the lowest total energy (E_{tot}). Cluster 2 was composed of 7 TPs that were characterized by the highest dipole moment (μ), energy of the lowest unoccupied molecular orbital (E_{LUMO}) and the lowest number of chlorine atoms. The final 4 TPs were grouped in cluster 3. Those molecules were characterized by the lowest MCIs, E_{HOMO}, MW, S and α but by the highest E_{tot}.

The different molecular properties of the TPs within each cluster also adumbrate deviating environmental behaviors, once more emphasizing the importance of the comprehensive knowledge of the existence of each occurring TP. This means that TPs which were classified into the same cluster as their parent molecule TCZ (including newly detected and yet

298 unknown TPs) may have similar behavior in the environment compared to TCZ. TPs which
299 were clustered away from TCZ may have different environmental fate and ecotoxicological
300 impact compared to TCZ. This can be especially problematic if their formation is ignored and
301 thus only TCZ and known TPs are quantified and monitored in the environment. To evaluate
302 these hypotheses, TyPol was applied for a second time in order to categorize 116 various
303 well-known pesticides and 19 TPs available in the TyPol database by analysis of both
304 molecular descriptors and their known environmental parameters (half-life in soil (DT_{50}),
305 sorption coefficient (K_{OC}), vapor pressure (P_{vap}) and bioconcentration factor (BCF)). The
306 compounds formed 6 clusters and median values of the molecular descriptors for each cluster
307 were obtained (an example is given for K_{OC} in the Supporting Information, tables S3 and S4,
308 and figure S3). In order to obtain a preliminary idea of the potential environmental behavior
309 of the detected TCZ TPs, their molecular descriptors were compared to the median values of
310 the 6 clusters formed by clustering the 116 pesticides and 19 TPs. When the molecular
311 descriptors of the well-known compounds of one cluster were highly similar to those of a
312 TCZ TP cluster, it was assumed that environmental parameters could be in a similar range.
313 TPs that were in the same cluster as their parent compound TCZ (cluster 1) were found to
314 overlap with a cluster of pesticides characterized by i) low P_{vap} (median P_{vap} = 0.026 mPa),
315 indicating a low volatility from soil and plants and thus low risk for short or long range
316 transfer through the atmosphere,[43] ii) low mobility in soil (median K_{OC} = 825 L/kg)
317 suggesting a low risk for groundwater contamination,[44] iii) low persistence in soil
318 (median DT_{50} = 37 days) and iv) low potential for bioaccumulation (median BCF = 78)
319 (Supporting Information, table S2). In the same manner, TCZ TPs forming cluster 2 had
320 molecular descriptor values similar to pesticides and TPs exhibiting i) a higher soil mobility
321 than TCZ (median K_{OC} = 217 L/kg), thus posing a moderate potential risk for groundwater
322 contamination,[44] ii) a higher potential for volatilization losses (median P_{vap} = 0.098 mPa) than

323 TCZ, iii) lower persistence in soil (median DT_{50} = 20 days) and iv) a lower potential for
324 bioaccumulation (median BCF = 44). Finally, TCZ TPs forming cluster 3 were the most
325 mobile molecules among the identified TPs (median K_{OC} of 123 L/kg) suggesting a
326 potentially higher risk for groundwater contamination. Furthermore, the TPs of cluster 3
327 posed the highest risk for volatilization losses (median P_{vap} = 1.600 mPa) suggesting a
328 significant risk for long or short range atmospheric transfer. Their persistence in soil was
329 assumed to be similar to that of cluster 2 (DT_{50} = 20 days) and lower than that of TCZ (cluster
330 1) and their potential for bioaccumulation was estimated to be low (BCF = 1) in comparison
331 to cluster 1 (BCF = 78) and cluster 2 (BCF = 44). These results are in agreement with
332 previous observations showing that when E_{HOMO}, MCIs, and MW decrease, sorption of
333 organic compounds tend to decrease, and when α, μ and MW increase, P_{vap} tend to decrease.[45]
334 Overall, the *in silico* approach using TyPol suggests that 23 of the 34 TCZ TPs targeted in this
335 study are expected to have a similar environmental behaviour as the parent compound. On the
336 other hand, the remaining 11 TPs may be more mobile and volatile than TCZ suggesting a
337 higher risk for groundwater and air contamination. However, their lower persistence and
338 lower bioaccumulation potential than TCZ may counterbalance this risk. This combined
339 method allows a first assessment of the fate of theoretical TPs in the environment identifying
340 those of potential concern that need to be intensively studied in a refined risk assessment
341 approach. If we consider TCZ as a case study, certain TPs from each cluster could be selected
342 based on their molecular features (e.g. selecting between some constitutional isomers, keto-
343 enol tautomers or others) and further considered for a refined ERA.
344 To our knowledge, our approach is the first which combines (i) comprehensive screening of
345 environmental samples (particularly soils) for suspected pesticide TPs based on a library of
346 empirical and theoretical TPs, (ii) detection of TPs without the need of analytical standards,
347 (iii) proposition of their molecular structures without the need of NMR spectroscopy and then

(iv) categorization of the detected TPs according to their physicochemical properties aiming to assess their environmental behavior. The approach described in this study could be applied for refining pesticide risk assessment and thus enable regulators and agrochemical industries to identify potential pesticide TPs showing relevant properties for further environmental and ecotoxicological assessment. Once the selected TPs are synthesized, absolute confirmation and quantification can be enabled and thus TP fate, environmental concentration and toxicity to ecosystems and human health can be assessed providing a refined and comprehensive risk assessment of a given pesticide.

ASSOCIATED CONTENT

Supporting information is available regarding the TCZ TP library, TP extraction from soil, the analytical method, chromatogram and MSC examples, the molecular typology method, detailed TyPol results for TCZ and its TPs and for the 116 pesticides and 19 TPs from the TyPol database. This material is available free of charge via the Internet.

AUTHOR INFORMATION

Corresponding authors

fmartin@dijon.inra.fr; luigi.lucini@unicatt.it

ACKNOWLEDGMENTS

The Ph.D. studies of Veronika Storck performed at the Doctoral School of Environment and Health (ED E2S) at the University of Burgundy were funded by the French Ministry of Education and Research (MESR). This work was done within the framework of the FP7-PEOPLE-2012-IAPP (Industry-Academia Partnerships and Pathways) Marie-Curie project 'Love-to-Hate' funded by the European Commission (Grant Agreement number 324349).

REFERENCES

(1) Hutzinger, O., Ed. *The handbook of environmental chemistry – anthropogenic compounds*, 3rd, ed.; Springer, Berlin Heidelberg, 2013.

(2) Lewis, S.L.; Maslin, M.A. Defining the Anthropocene. *Nature* **2015**, 519, 171-180.

(3) *Pesticides industry sales and usage - 2006 and 2007 market estimates*; U.S. EPA, 2011; http://www.epa.gov/opp00001/pestsales/07pestsales/market_estimates2007.pdf.

(4) Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/ 117/EEC and 91/414/EEC. Regulation EC 1107/2009, 2009; http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009R1107&from=en.

(5) *Technical overview of ecological risk assessment. Analysis phase: exposure characterization – fate and transport of pesticides*; U.S. EPA, 2015; http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_exp.htm.

(6) Bozdogan, A.M. Assessment of total risk on non-target organisms in fungicide application for agricultural sustainability. *Sustainability* **2014**, 6, 1046-1058.

(7) *A desk study on pesticide metabolites, degradation and reaction products to inform the inspectorate's position on monitoring requirements. Final report for drinking water inspectorate*; Sinclair, C.; Van Beinum, W.; Adams, C.; Bevan, R.; Levy, L.; Parsons, S.; Goslan, E.; Baumann, G.; The Food and Environment Research Agency **2010**; http://www.google.fr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCEQFjAA&url=http%3A%2F%2Frandd.defra.gov.uk%2FDocument.aspx%3FDocument%3DWT1221_9026_FRP.pdf&ei=Bj6aVaHVFIG-

UOOzgqgF&usg=AFQjCNEgPz56VUkW1sg01QHQXJx18kKCWg&bvm=bv.96952980,d.d
24.

(8) Tixier, C.; Bogaerts, P.; Sancelme, M.; Bonnemoy, F.; Twagilimana, L.; Cuer, A.; Bohatier, J.; Veschambre, H. Fungal biodegradation of a phenylurea herbicide, diuron: structure and toxicity of metabolites. *Pest Management Science* **2000**, 56, 455-462.

(9) Boxall, A.B.A.; Sinclair, C.J.; Fenner, K.; Kolpin, D.; Maund, S.J. Peer reviewed: When synthetic chemicals degrade in the environment. *Environmental Science & Technology* **2004**, 38 (19), 368A-375A.

(10) Fenner, K.; Canonica, S.; Wackett, L.P.; Elsner, M. Evaluating pesticide degradation in the environment: Blind spots and emerging opportunities. *Science* **2013**, 341, 752-758.

(11) Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for community action in the field of water policy. EU Water Framework Directive 2000/60/EC, 2000; http://eur-lex.europa.eu/resource.html?uri=cellar:5c835afb-2ec6-4577-bdf8-756d3d694eeb.0004.02/DOC_1&format=PDF.

(12) Directive 2008/50/EC of the European Parliament and of the Council on ambient air quality and cleaner air for Europe. EU Air Quality Directive 2008/50/EC, 2008; http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:152:0001:0044:EN:PDF.

(13) Proposal for a directive of the European Parliament and of the Council establishing a framework for the protection of soil and amending Directive 2004/35/EC. Proposal for

Directive 2004/35/EC, 2006; http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52006PC0232&from=EN.

(14) *Guidance document on pesticide residue analytical methods*; OECD Guideline ENV/JM/MONO(2007)17, 2007; http://www.oecd.org/env/ehs/pesticides-biocides/publicationsonpesticideresidues.htm.

(15) Ibanez, M.; Sancho, J.V.; Pozo, O.J.; Niessen, W.; Hernandez, F. Use of quadrupole time-of-flight mass spectrometry in the elucidation of unknown compounds present in environmental water. *Rapid Communications in Mass Spectrometry* **2005**, 19, 169-178.

(16) Reitzel, L.A.; Dalsgaard, P.W.; Muller, I.B.; Cornett, C. Identification of ten new designer drugs by GC-MS, UPLC-QTOF-MS, and NMR as part of a police investigation of a Danish internet company. *Drug Testing and Analysis* **2012**, 4 (5), 342-254.

(17) Zeng, W.; Han, H.; Tao, Y.; Yang, L.; Wang, Z.; Chen, K. Identification of bio-active metabolites of gentiopicroside by UPLC/Q-TOF MS and NMR. *Biomedical Chromatography* **2013**, 27, 1129-1136.

(18) Servien, R.; Mamy, L.; Li, Z.; Rossard, V.; Latrille, E.; Bessac, F.; Patureau, D.; Benoit, P. TyPol – A new methodology for organic compounds clustering based on their molecular characteristics and environmental behavior. *Chemosphere* **2014**, 111, 613-622.

(19) Börner, H.; Schlüter, K.; Aumann, J. *Pflanzenkrankheiten und Pflanzenschutz*. Springer, Berlin Heidelberg, 2009.

(20) *Conclusion on the peer review of the pesticide risk assessment of the active substance tebuconazole*; EFSA, 2014; http://www.efsa.europa.eu/fr/efsajournal/doc/3485.pdf.

(21) Rieke, S.; Koehn, S.; Hirsch-Ernst, K.; Pfeil, R.; Kneuer, C.; Marx-Stoelting, P. Combination effects of (tri)azole fungicides on hormone production and xenobiotic metabolism in a human placental cell line. *International Journal of Environmental Research and Public Health* **2014**, 11, 9660-9679.

(22) Shalini, K.; Kumar, N.; Drabu, S.; Kumar Sharma, P. Advances in synthetic approach to and antifungal activity of triazoles. *Beilstein Journal of Organic Chemistry* **2011**, 7, 668-677.

(23) *Folicur® EW250 – directions for use*; Bayer Crop Science 2011; http://www.bayercropscience.ie/labels/Folicur_250_EW.pdf.

(24) *Soil quality – Sampling – Part 6: Guidance on the collection, handling and storage of soil under aerobic conditions for the assessment of microbiological processes, biomass and diversity in the laboratory*; ISO 10381-6, 2009; https://www.iso.org/obp/ui/#iso:std:iso:10381:-6:en.

(25) Papadopoulou, E.S.; Karas, P.A.; Nikolaki, S.; Storck, V.; Ferrarini, A.; Fornasier, F.; Ferrari, F.; Trevisan, M.; Martin-Laurent, F.; Tsiamis, G.; Karpouzas, D.G. A lab-to-field experimental approach to study the dissipation, metabolism and soil microbial ecotoxicity of isoproturon, tebuconazole and chlorpyrifos. *IUPAC International Congress of Pesticide Chemistry, San Francisco* **2014**.

(26) Hill, A. W.; Mortishire-Smith R. J. Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. *Rapid Communications in Mass Spectrometry* **2005**, 19, 3111-3118.

(27) Sehnem, N.; Souza-Cruz, P.; Peralba, M.; Zachia Ayub, M. Biodegradation of tebuconazole by bacteria isolated from contaminated soils. Journal *of Environmental Science and Health Part B* **2010**, 45, 67-72.

(28) Woo, C.; Daniels, B.; Stirling, R.; Morris, P. Tebuconazole and propiconazole tolerance and possible degradation by Basidiomycetes: A wood-based bioassay. *International Biodeterioration and Biodegradation* **2010**, 64, 403-408.

(29) Obanda, D.N.; Shupe, T.F. Biotransformation of tebuconazole by microorganisms: Evidence of a common mechanism. *Wood and Fiber Science* **2009**, 41 (2), 157-167.

(30) Calza, P.; Baudino, S.; Aigotti, R.; Baiocchi, C.; Branca, P.; Pelizzetti, E. High-performance liquid chromatographic/tandem mass spectrometric identification of the phototransformation products of tebuconazole on titanium dioxide. *Journal of Mass Spectrometry* **2002**, 37, 566-576.

(31) Herrero-Hernandez, E.; Andrades, M. S.; Marin-Benito, J. M.; Sanchez-Martin, M. J.; Rodriguez-Cruz, M.S. Field-scale dissipation of tebuconazole in a vineyard soil amended with spent mushroom substrate and its potential environmental impact. *Ecotoxicology and Environmental Safety* **2011**, 74, 1480-1488.

(32) Potter, T.L.; Strickland, T.C.; Joo, H.; Culbreath, A.K. Accelerated soil dissipation of tebuconazole following multiple applications to peanut. *Journal of Environmental Quality* **2005**, 34, 1205-1213.

(33) Strickland, T.C.; Potter, T.L.; Hyun, J. Tebuconazole dissipation and metabolism in Tifton loamy sand during laboratory incubation. *Pest Management Science* **2004**, 60, 703-709.

(34) White, P.M.; Potter, T.L.; Culbreath, A.K. Fungicide dissipation and impact on metolachlor aerobic soil degradation and soil microbial dynamics. *Science of the Total Environment* **2010**, 408, 1393-1402.

(35) Schermerhorn, P.G.; Golden, P.E.; Krynitsky, A.J.; Leimkuehler, W.M. Determination of 22 triazole copounds including parent fungicides and metabolites in apples, peaches, flour and water by liquid chromatography/tandem mass spectrometry. *Journal of AOAC International* **2005**, 88 (5), 1491-1502.

(36) *Scientific Report 176 - Conclusion on the peer review of tebuconazole*; EFSA, 2008; http://www.efsa.europa.eu/en/scdocs/doc/176r.pdf.

(37) *Pesticide Evaluation Tebuconazole 188*; FAO, 1994; http://www.fao.org/fileadmin/templates/agphome/documents/Pests_Pesticides/JMPR/Evaluation94/tebucona.pdf.

(38) *Tebuconazole Memorandum*; U.S. EPA, 1990; http://www.epa.gov/pesticides/chem_search/cleared_reviews/csr_PC-128997_18-Sep-90_025.pdf.

(39) Schymanski, E.L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H.P.; Hollender, J. Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environmental Science and Technology* **2014**, 48, 2097-2098.

(40) Giacomazzi, S.; Cochet, N. Environmental impact of diuron transformation: a review. *Chemosphere* **2004**, 56, 1021-1032.

(41) Horn, W.; Jann, O.; Wilke, O. Suitability of small environmental chambers to test the emission of biocides from treated materials into the air. *Atmospheric Environment* **2003**, 37, 5477-5483.

(42) Wang, X.; Wang, X.; Zhang, H.; Wu, C.; Wang, X.; Xu, H.; Wang, X.; Li, Z. Enantioselective degradation of tebuconazole in cabbage, cucumber and soils. *Chirality* **2012**, 24, 104-111.

(43) *Pesticides in air: Considerations for exposure assessment. Report of the FOCUS Working Group on Pesticides in Air*; EC Document Reference SANCO/10553/2006, FOCUS, 2008; http://focus.jrc.ec.europa.eu/ai/docs/FOCUS_AIR_GROUP_REPORT-FINAL.pdf.

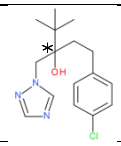
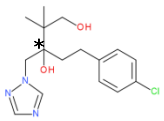
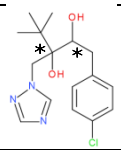
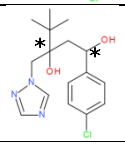
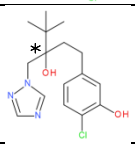
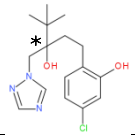
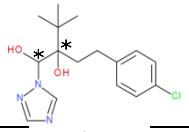
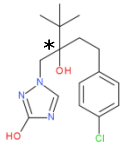
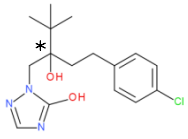
546 (44) McCall, P.J.; Swann, R.L.; Laskowski, D.A.; Unger, S.M.; Vrona, S.A.; Dishburger, H.J.
547 Estimation of chemical mobility in soil from liquid chromatographic retention times. *Bulletin*
548 *of Environmental Contamination and Toxicology* **1980**, 24, 190-195.
549
550 (45) Mamy, L.; Patureau, D.; Barriuso, E.; Bedos, C.; Bessac, F.; Louchart, X.; Martin-
551 Laurent, F.; Miège, C.; Benoit, P. Prediction of the fate of organic compounds in the
552 environment from their molecular properties: A review. *Critical Reviews in Environmental*
553 *Science and Technology* **2015**, 45, 1277-1377.
554

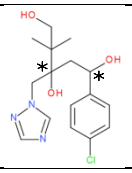
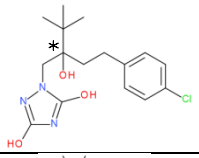
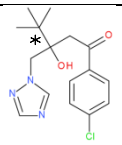
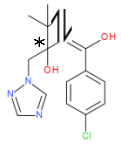
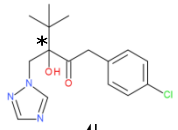
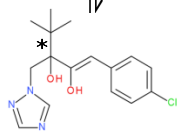
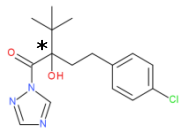
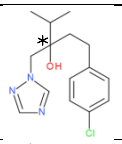
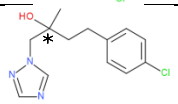
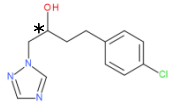
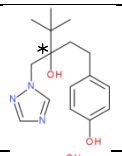
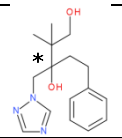
555 Abstract art

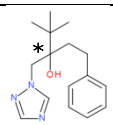
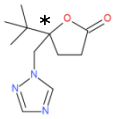
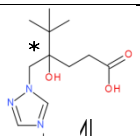
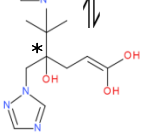
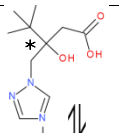
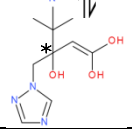
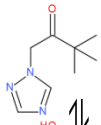
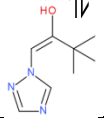
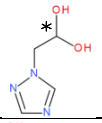
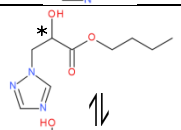
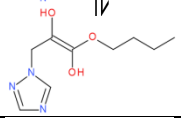
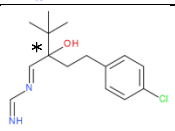
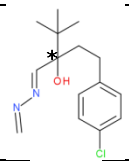
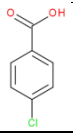
556 Picture Tractor spraying wheat field: © Dusan Kostic, Fotolia Credit Licence 82893177; Picture QTOF-MS: © Agilent Technologies, Inc. 2015. Reproduced with Permission, Courtesy of Agilent Technologies, Inc.

557

Table 1: TCZ TPs detected in this study by suspect screening analysis. Reference to the degree of confidence on identification (a. detected by MS (orange hooklet); b. selected by target tandem MS (yellow hooklet); c. proposed molecular structure by MSC (green hooklet)), name, structure, mass and reference if empirical of each TP are provided. The time points at which a TP was detected and the cluster into which it was categorized are given. Blue boxes group TPs having the same mass.

compound name	molecular structure	mass [g/mol]	reference	detected in this study	RT [min]	details	TyPol cluster
TCZ, 1-(4-chlorophenyl)-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentan-3-ol		307.15	38	✓ a,b,c	9.38	detected at 0,3,7,21,35, 56,70,105 and 125 d	1
TP 1, 5-(4-chlorophenyl)-2,2-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentane-1,3-diol		323.14	29	✓ a	9.24	detected at 21,35 and 70 d	1
TP 2, 1-(4-chlorophenyl)-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentane-2,3-diol		323.14	37	✓ a	9.24	detected at 21,35 and 70 d	1
TP 3, 1-(4-chlorophenyl)-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentane-1,3-diol		323.14	theoretical TP	✓ a	9.24	detected at 21,35 and 70 d	1
TP 4, 2-chloro-5-[3-hydroxy-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentyl]phenol		323.14	theoretical TP	✓ a	9.24	detected at 21,35 and 70 d	1
TP 5, 5-chloro-2-[3-hydroxy-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentyl]phenol		323.14	37	✓ a	9.24	detected at 21,35 and 70 d	1
TP 6, 2-[2-(4-chlorophenyl)ethyl]-3,3-dimethyl-1-(1,2,4-triazol-1-yl)butane-1,2-diol		323.14	theoretical TP	✓ a	9.24	detected at 21,35 and 70 d	1
TP 7, 1-[2-[2-(4-chlorophenyl)ethyl]-2-hydroxy-3,3-dimethyl-butyl]-1,2,4-triazol-3-ol		323.14	theoretical TP	✓ a	9.24	detected at 21,35 and 70 d	1
TP 8, 2-[2-[2-(4-chlorophenyl)ethyl]-2-hydroxy-3,3-dimethyl-butyl]-1,2,4-triazol-3-ol		323.14	theoretical TP	✓ a	9.24	detected at 21,35 and 70 d	1

TP 9, 1-(4-chlorophenyl)-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentane-1,3,5-triol		339.14	37	✓	a	9.06	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 10, 1-[2-[2-(4-chlorophenyl)ethyl]-2-hydroxy-3,3-dimethyl-butyl]-1,2,4-triazole-3,5-diol		339.14	theoretical TP	✓	a	9.06	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 11, 1-(4-chlorophenyl)-3-hydroxy-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentan-1-one		321.12	30-33	✓	a,b,c	9.41	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 12, (E)-1-(4-chlorophenyl)-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pent-1-ene-1,3-diol		321.12	37	✓	a,b,c	9.41	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 13, 1-(4-chlorophenyl)-3-hydroxy-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentan-2-one		321.12	theoretical TP	✓	a,b,c	9.41	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 14, (Z)-1-(4-chlorophenyl)-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pent-1-ene-2,3-diol		321.12	theoretical TP	✓	a,b,c	9.41	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 15, 2-[2-(4-chlorophenyl)ethyl]-2-hydroxy-3,3-dimethyl-1-(1,2,4-triazol-1-yl)butan-1-one		321.12	theoretical TP	✓	a,b,c	9.41	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 23, 1-(4-chlorophenyl)-4-methyl-3-(1,2,4-triazol-1-ylmethyl)pentan-3-ol		293.13	37	✓	a,b,c	10.15	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 24, 4-(4-chlorophenyl)-2-methyl-1-(1,2,4-triazol-1-yl)butan-2-ol		265.10	37	✓	a	6.54	detected at 105 and 125 d	1
TP 25, 4-(4-chlorophenyl)-1-(1,2,4-triazol-1-yl)butan-2-ol		251.08	37	✓	a	4.24	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 29, 4-[3-hydroxy-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentyl]phenol		289.18	theoretical TP	✓	a,b	8.20	detected at 105 and 125 d	1
TP 30, 2,2-dimethyl-5-phenyl-3-(1,2,4-triazol-1-ylmethyl)pentane-1,3-diol		289.18	theoretical TP	✓	a,b	8.20	detected at 105 and 125 d	1

TP 31, 4,4-dimethyl-1-phenyl-3-(1,2,4-triazol-1-ylmethyl)pentan-3-ol		273.18	theoretical TP	✓	a,b,c	9.65	detected at 3,7,21,35, 56,70,105 and 125 d	1
TP 32, 5-tert-butyl-5-(1,2,4-triazol-1-ylmethyl)tetrahydrofuran-2-one		223.13	29,34	✓	a,b,c	5.13	detected at 35,56,70, 105 and 125 d	2
TP 33, 4-hydroxy-5,5-dimethyl-4-(1,2,4-triazol-1-ylmethyl)hexanoic acid		241.14	33	✓	a	8.38	detected at 3,7,21,35 and 56 d	2
TP 34, 5,5-dimethyl-4-(1,2,4-triazol-1-ylmethyl)hex-1-ene-1,1,4-triol		241.14	33	✓	a	8.38	detected at 3,7,21,35 and 56 d	2
TP 35, 3-hydroxy-4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pentanoic acid		227.13	32	✓	a	4.72	detected at 3,7,21,35 and 56 d	2
TP 36, 4,4-dimethyl-3-(1,2,4-triazol-1-ylmethyl)pent-1-ene-1,1,3-triol		227.13	32	✓	a	4.72	detected at 3,7,21,35 and 56 d	2
TP 37, 3,3-dimethyl-1-(1,2,4-triazol-1-yl)butan-2-one		167.11	33	✓	a,b,c	9.28	detected at 35,56,70, 105 and 125 d	3
TP 38, (E)-3,3-dimethyl-1-(1,2,4-triazol-1-yl)but-1-en-2-ol		167.11	33	✓	a,b,c	9.28	detected at 35,56,70, 105 and 125 d	3
TP 46, 2-(1,2,4-triazol-1-yl)ethane-1,1-diol		129.05	35,36	✓	a,b	0.68	detected at 56,70,105 and 125 d	3
TP 49, butyl 2-hydroxy-3-(1,2,4-triazol-1-yl)propanoate		213.11	38	✓	a	7.27	detected at 3,7,21,35 and 56 d	2
TP 50, (Z)-1-butoxy-3-(1,2,4-triazol-1-yl)prop-1-ene-1,2-diol		213.11	38	✓	a	7.27	detected at 3,7,21,35 and 56 d	2
TP 66, (NE)-N-[2-[2-(4-chlorophenyl)ethyl]-2-hydroxy-3,3-dimethyl-butylidene]formamidine		280.13	29	✓	a,b,c	8.14	detected at 7,21,35,56,70 and 105 d	1
TP 67, 1-(4-chlorophenyl)-4,4-dimethyl-3-[(E)-(methylenehydrazono)methyl] pentan-3-ol		280.13	29	✓	a,b,c	8.14	detected at 7,21,35,56,70 and 105 d	1
TP 75, 4-chlorobenzoic acid		156.00	37	✓	a	9.05	detected at 70 and 105 d	3

* Chiral center

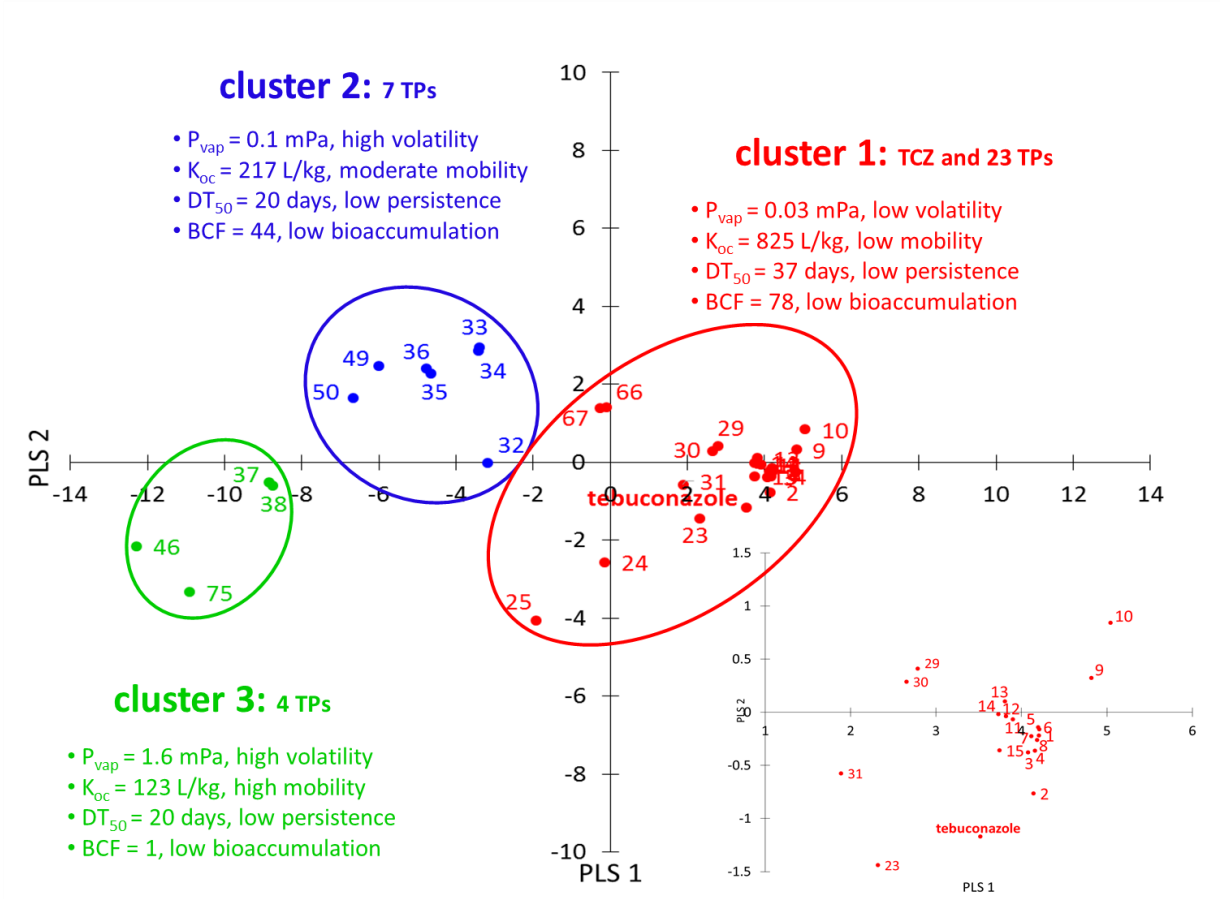


Figure 1: Categorization of TCZ and its 34 detected TPs into 3 clusters according to their molecular descriptors on the two main components of the PLS regression (PLS 1 and PLS 2) and their estimated environmental parameters. Formed clusters are defined by a different color code. Each TP has its respective number. The inserted graph serves as enlargement of a part of cluster 1.